

ระบบฐานข้อมูล

Chapter 03: Entity, Relation and Normalization

อ.นุชนาถ สัตย์วินิจ

ภาควิชาวิศวกรรมคอมพิวเตอร์ วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์

Topics



ER Model to Relational Database



Normalization



การแปลงแบบจำลองอีอาร์ให้เป็นฐานข้อมูลเชิงสัมพันธ์

- 1. แบบชนิดเอนทิตีปรกติ
- 2. แบบชนิดเอนทิตีอ่อน
- 3. ความสัมพันธ์แบบหนึ่งต่อหนึ่ง
- 4. ความสัมพันธ์แบบหนึ่งต่อกลุ่ม
- 5. ความสัมพันธ์แบบกลุ่มต่อกลุ่ม
- 6. แอตทริบิวต์หลายค่า
- 7. ความสัมพันธ์ตั้งแต่สองเอนทิตีขึ้นไป



ขั้นตอนที่ 1: แบบชนิดเอนทิตีปรกติ

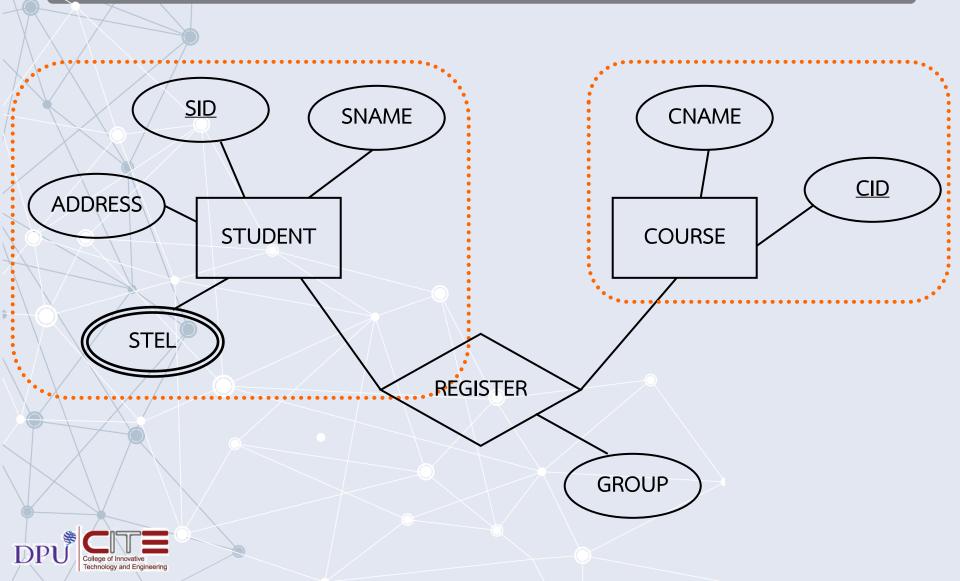
แปลงแบบชนิดเอนทิตีปรกติเป็นหนึ่งตารางความสัมพันธ์

• แอตทริบิวต์ของความสัมพันธ์มาจากแอตทริบิวต์ของแบบชนิดเอนทิตี ยกเว้น แอตทริบิวต์หลายค่า

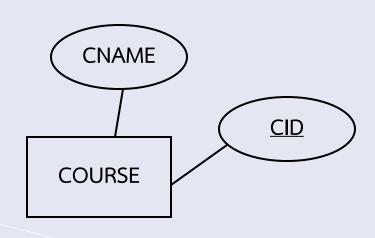
• กุญแจหลักของตารางความสัมพันธ์คือแอตทริบิวต์หลักของแบบชนิด เอนทิตี



ตัวอย่าง: แบบชนิดเอนทิตีปรกติ



ตัวอย่าง: แบบชนิดเอนทิตีปรกติ

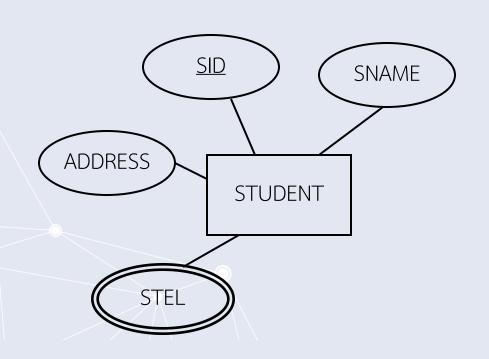


COURSE





ตัวอย่าง: แบบชนิดเอนทิตีปรกติ





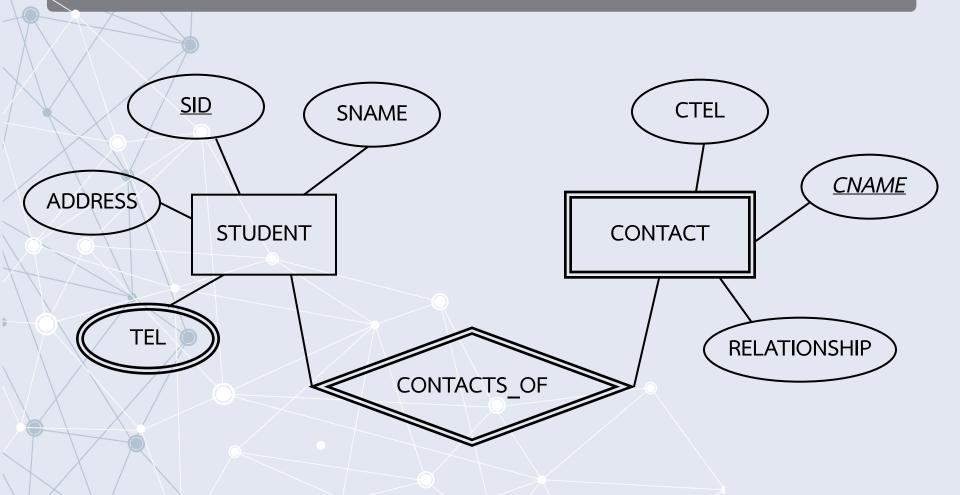
ขั้นตอนที่ 2: แบบชนิดเอนทิตีอ่อน

แปลงแบบชนิดเอนทิตีอ่อนเป็นหนึ่งตารางความสัมพันธ์

 กุญแจหลักของตารางความสัมพันธ์ ประกอบด้วยกุญแจย่อยของแบบ ชนิดเอนทิตีอ่อนนั้น รวมกับแอตทริบิวต์หลักของแบบชนิดเอนทิตีแม่

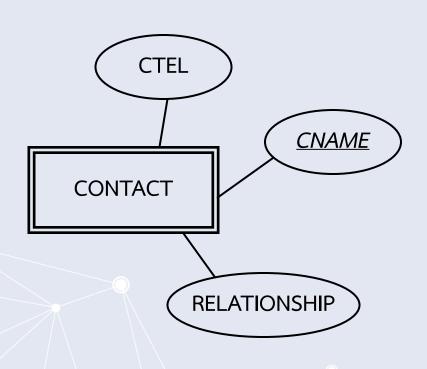


ตัวอย่าง: แบบชนิดเอนทิตีอ่อน





ตัวอย่าง: แบบชนิดเอนทิตีอ่อน



CONTACT

SID CNAME CTEL RELATIONSHIP	
-----------------------------	--



<u>ขั้นตอนที่ 3: ความสัมพันธ์แบบหนึ่งต่อหนึ่ง</u>

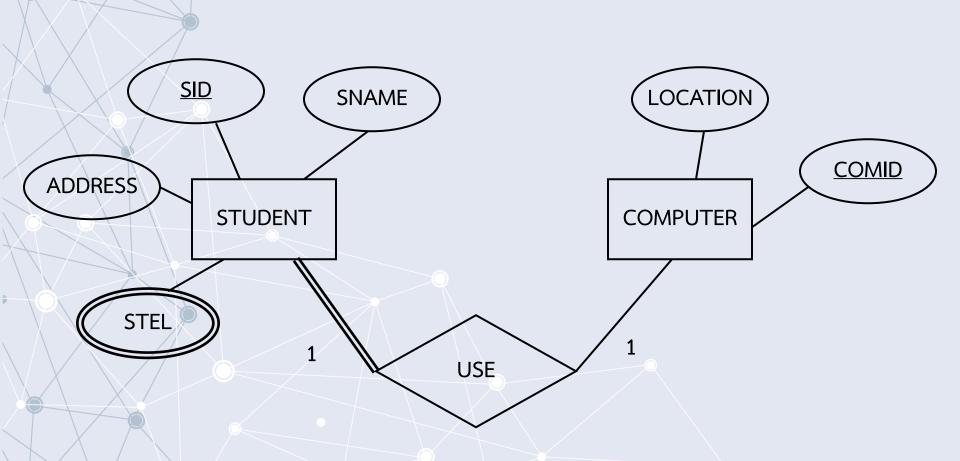
ไม่ต้องสร้างตารางความสัมพันธ์ใหม่

• ให้ฝั่งที่มีส่วนร่วมทุกส่วน (total participation) เป็นหลัก

• นำแอตทริบิวต์หลักของแบบชนิดเอนทิตีฝั่งที่มีส่วนร่วมบางส่วน (partial participation) มาเป็นกุญแจนอก

• เติมแอตทริบิวต์ของความสัมพันธ์





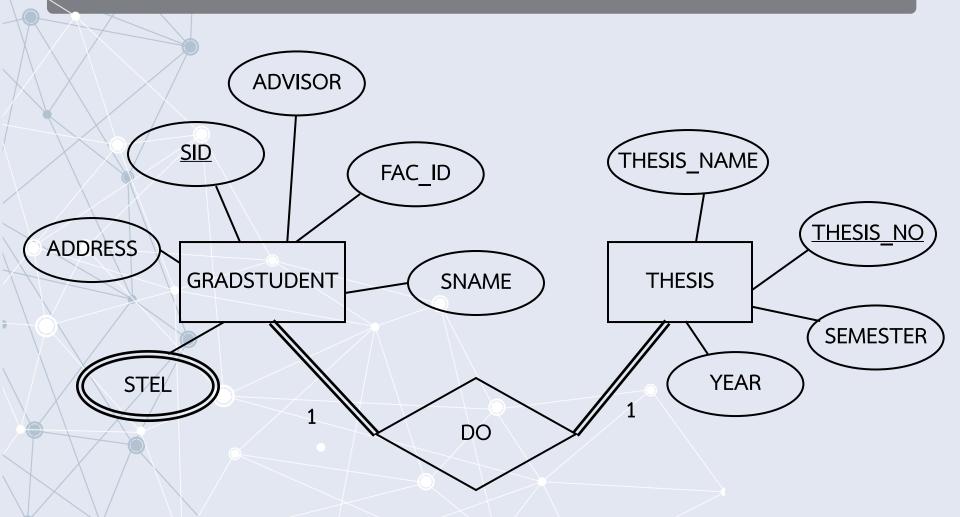


• จากขั้นตอนที่ 1



• ขั้นตอนที่ 3







ขั้นตอนที่ 4: ความสัมพันธ์แบบหนึ่งต่อกลุ่ม

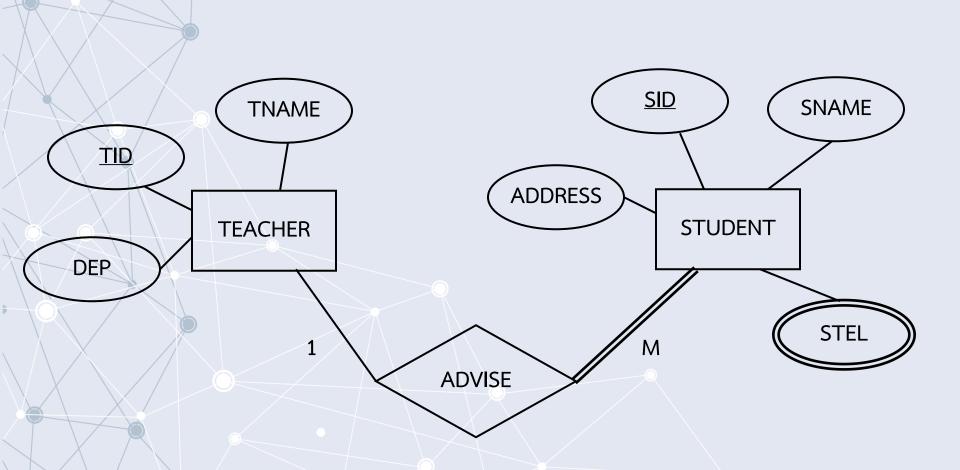
ไม่ต้องสร้างตารางความสัมพันธ์ใหม่

• ให้ตารางความสัมพันธ์ฝั่งกลุ่ม (many) เป็นหลัก

• นำแอตทริบิวต์หลักของแบบชนิดเอนทิตีฝั่งหนึ่ง (one) มาเป็นกุญแจ นอก

• เติมแอตทริบิวต์ของความสัมพันธ์







• จากขั้นตอนที่ 1



• ขั้นตอนที่ 4



ขั้นตอนที่ 5: ความสัมพันธ์แบบกลุ่มต่อกลุ่ม

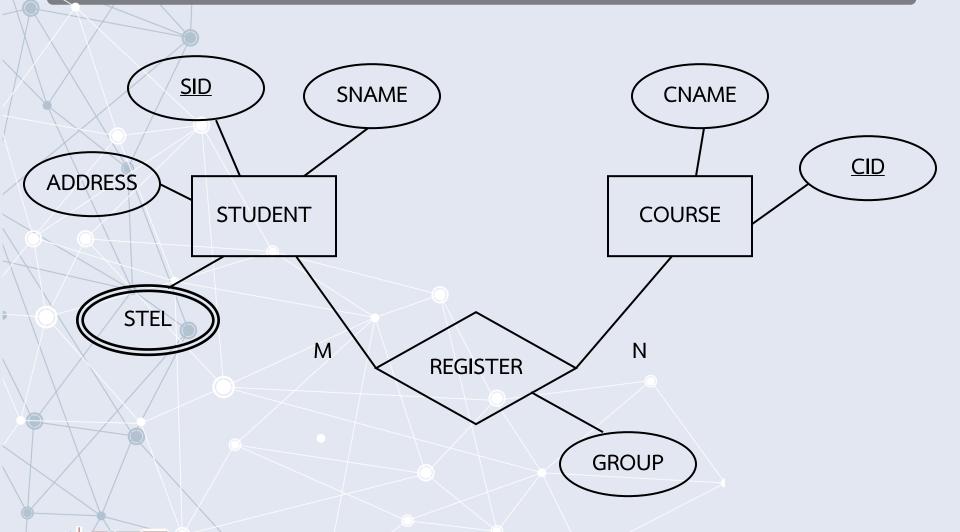
• แปลงความสัมพันธ์แบบกลุ่มต่อกลุ่มเป็นหนึ่งตารางความสัมพันธ์

กุญแจหลักของตารางความสัมพันธ์ประกอบด้วยแอตทริบิวต์หลักของ
 แบบชนิดเอนทิตีที่เกี่ยวข้อง

แอตทริบิวต์ของตารางความสัมพันธ์มาจากแอตทริบิวต์ของ
 ความสัมพันธ์



ตัวอย่าง: ความสัมพันธ์แบบกลุ่มต่อกลุ่ม





ตัวอย่าง: ความสัมพันธ์แบบกลุ่มต่อกลุ่ม

• จากขั้นตอนที่ 1

• ขั้นตอนที่ 5



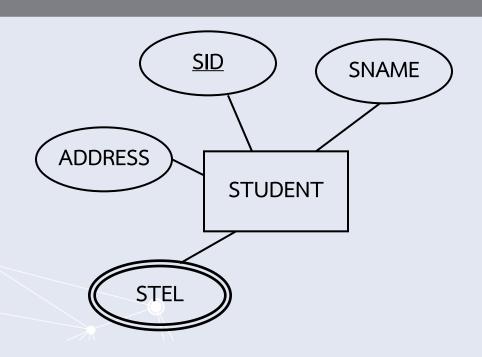
ขั้นตอนที่ 6: แอตทริบิวต์หลายค่า

แปลงแอตทริบิวต์หลายค่าเป็นหนึ่งตารางความสัมพันธ์

• กุญแจหลักของตารางความสัมพันธ์มาจากแอตทริบิวต์หลายค่านั้นรวม กับแอตทริบิวต์หลักของแบบชนิดเอนทิตี



ตัวอย่าง: แอตทริบิวต์หลายค่า





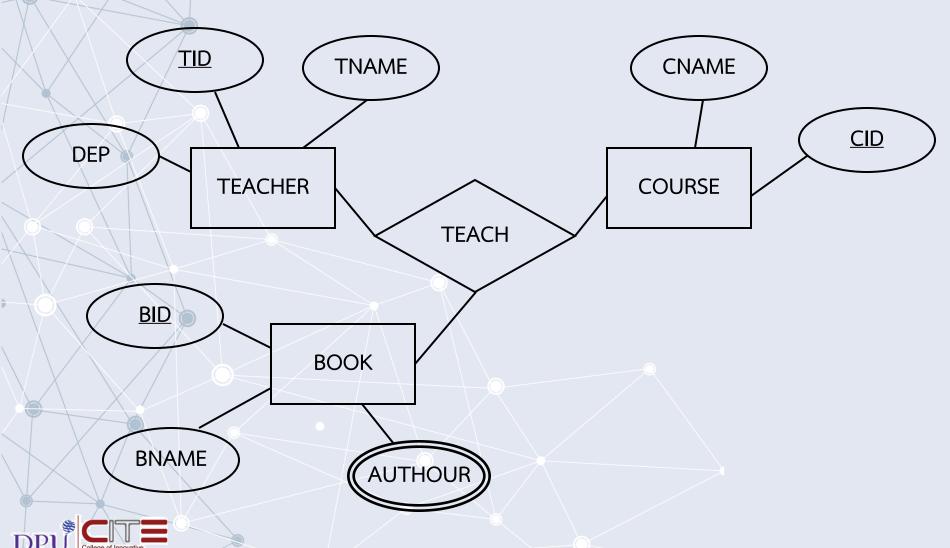
ขั้นตอนที่ 7: ความสัมพันธ์ตั้งแต่สองเอนทิตีขึ้นไป

• ความสัมพันธ์ที่เกี่ยวข้องกับแบบชนิดเอนทิตีตั้งแต่สองตัวขึ้นไป (N-ary relationship) ให้แปลงเป็นหนึ่งตารางความสัมพันธ์

• กุญแจหลักของตารางความสัมพันธ์ประกอบด้วยแอตทริบิวต์หลักของ แบบชนิดเอนทิตีที่เกี่ยวข้อง



ตัวอย่าง: ความสัมพันธ์ตั้งแต่สองเอนทิตีขึ้นไป



• จากขั้นตอนที่ 1



ตัวอย่าง: ความสัมพันธ์ตั้งแต่สองเอนทิตีขึ้นไป

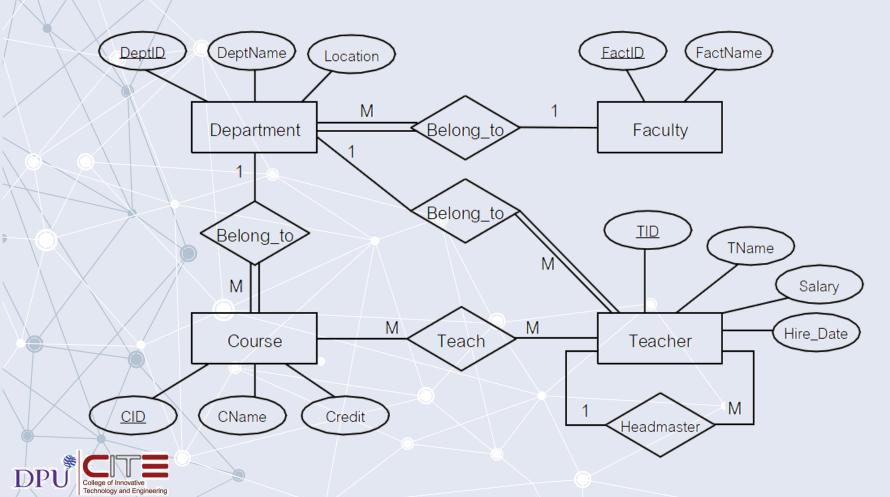
• จากขั้นตอนที่ 6

• ขั้นตอนที่ 7

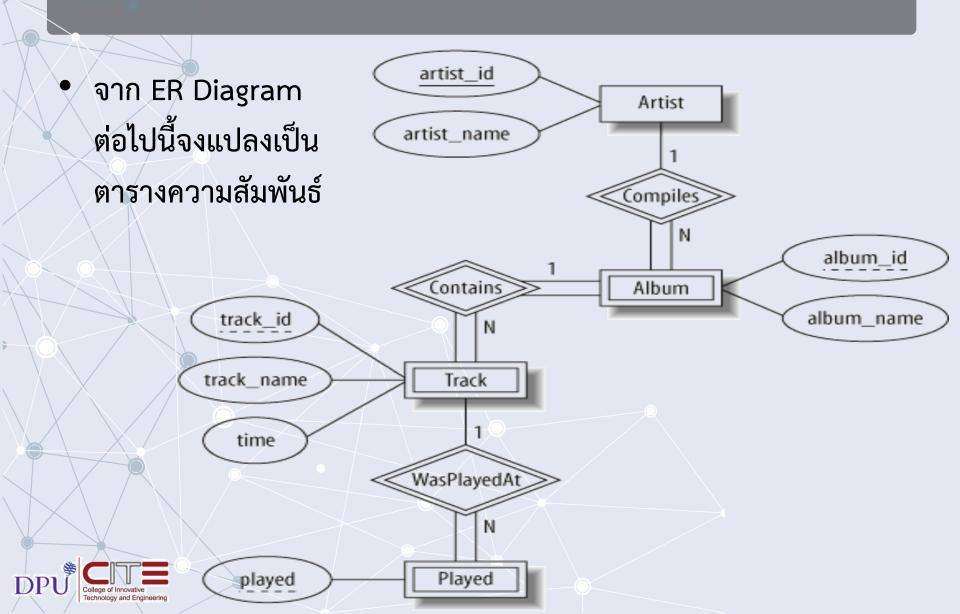


Problem 1

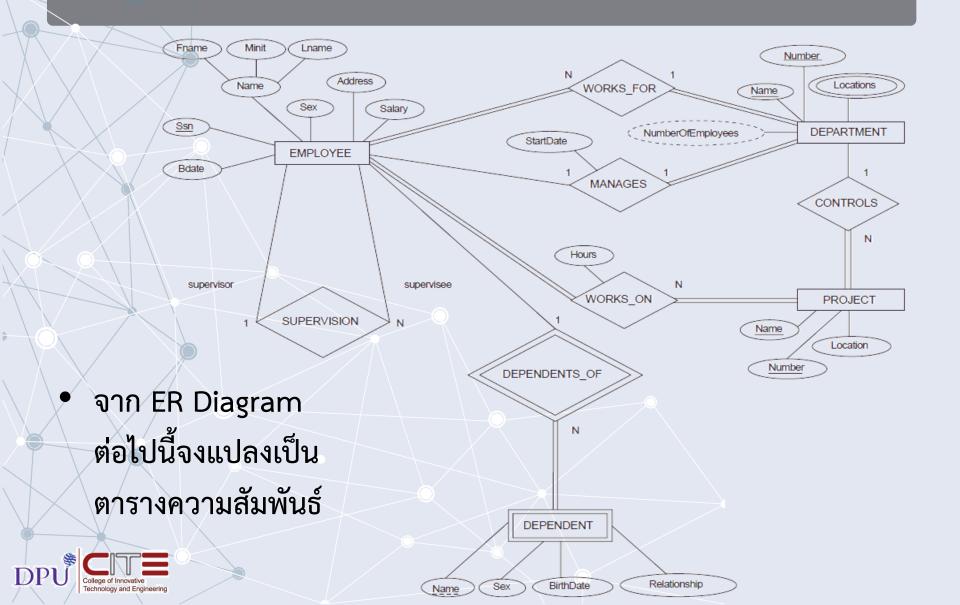
• จาก ER Diagram ต่อไปนี้จงแปลงเป็นตารางความสัมพันธ์



Problem 2



Problem 3



Normalization

• คือกระบวนการปรับเปลี่ยนโครงสร้างของ relation schema ให้ อยู่ในรูปแบบ "Normal Form" ในลำดับต่างๆ

- กำหนด Attribute ให้กับแต่ละตารางความสัมพันธ์ เพื่อให้ได้ โครงสร้างของตารางที่ดี
 - ลดความซ้ำซ้อนของข้อมูล
 - ลดความผิดปรกติของข้อมูล



ตัวอย่างการออกแบบฐานข้อมูลที่จะมีปัญหาของความซ้ำซ้อนตามมา

EMP_ID	EMP_NAME	POSITION	SALARY	BRANCH_ID	ADDRESS
AM021	พีรพล โพธิ์ทัย	MANAGER	30000	B001	BANGKOK
AM035	ศิริรัตน์ สายคุณากร	ASSISTANT MANAGER	20000	B005	RATCHABURI
AM010	เพชราภรณ์ เปรี่ยมศิริ	SECRETARY	20000	B004	CHIANGMAI
AM005	ธวัชชัย ด่านลาเคน	MANAGER	30000	B005	RATCHABURI



ตัวอย่างปัญหาความซ้ำซ้อนในข้อมูล

\	EMP_ID	EMP_NAME	POSITION	SALARY	BRANCH_ID	ADDRESS
1	AM021	พีรพล โพธิ์ทัย	MANAGER	30000	B001	BANGKOK
	AM035	ศิริรัตน์ สายคุณากร	ASSISTANT MANAGER	20000	B005	RATCHABURI
	AM010	เพชราภรณ์ เปรี่ยมศิริ	SECRETARY	20000	B004	CHIANGMAI
/	AM005	ธวัชชัย ด่านลาเคน	MANAGER	30000	B005	RATCHABURI

<u>ความผิดพลาดจากการเพิ่ม</u>

• ถ้าต้องการเพิ่มพนักงานใหม่ ที่อยู่สาขา B004



ตัวอย่างปัญหาความซ้ำซ้อนในข้อมูล

EMP_ID	EMP_NAME	POSITION	SALARY	BRANCH_ID	ADDRESS
AM021	พีรพล โพธิ์ทัย	MANAGER	30000	B001	BANGKOK
AM035	ศิริรัตน์ สายคุณากร	ASSISTANT MANAGER	20000	B005	RATCHABURI
AM010	เพชราภรณ์ เปรี่ยมศิริ	SECRETARY	20000	B004	CHIANGMAI
AM005	ธวัชชัย ด่านลาเคน	MANAGER	30000	B005	RATCHABURI

ความผิดพลาดจากการเพิ่ม

• ถ้าต้องการเพิ่มสาขา จะมีปัญหาคือ ตารางนี้มีทั้งข้อมูลพนักงานและข้อมูล สาขาอยู่รวมกัน



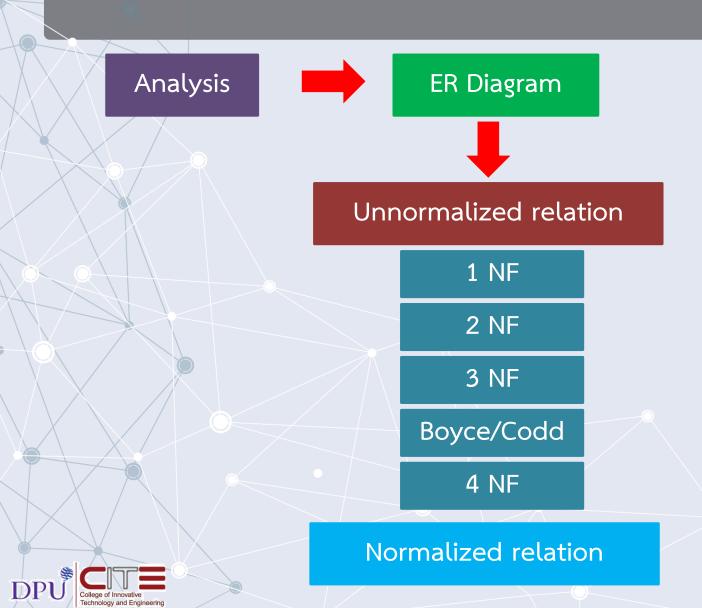
ตัวอย่างการออกแบบฐานข้อมูลที่ดี

EMP_ID	EMP_NAME	POSITION	SALARY	BRANCH_ID
AM021	พีรพล โพธิ์ทัย	MANAGER	30000	B001
AM035	ศิริรัตน์ สายคุณากร	ASSISTANT MANAGER	20000	B005
AM010	เพชราภรณ์ เปรี่ยมศิริ	SECRETARY	20000	B004
AM005	ธวัชชัย ด่านลาเคน	MANAGER	30000	B005

BRANCH_ID	ADDRESS
B001	BANGKOK
B005	RATCHABURI
B004	CHIANGMAI



Normalization



รูปแบบบรรทัดฐาน

- มีได้ถึง 5 ระดับ ในระดับที่ 3 ก็จัดว่าเพียงพอสำหรับการออกแบบฐานข้อมูล ในปัจจุบัน
 - 1 NF กำจัด repeating group (กำจัดกลุ่มของข้อมูลที่มีความซ้ำซ้อน)

— 2 NF - กำจัด partial dependency (กำจัดการขึ้นต่อกันบางส่วน)

— 3 NF - กำจัด transitive dependency (กำจัดการขึ้นต่อกันของ Attribute ที่ ไม่ใช่คีย์หลัก)



รูปแบบบรรทัดฐาน

1NF "ขจัด Repeating Group"

2NF "ขจั๊ด Partial Dependency"

3NF "ขจัด Transitive Dependency"

BCNF "ขจัดปัญหา Overlap Multiple Candidate Key"

4NF "ขจัด Nontrivial Multivalued Dependency"

5NF "ขจัด Join Dependency"



รูปแบบบรรทัดฐานที่หนึ่ง(First Normal Form: 1NF)

- ความสัมพันธ์ใด ๆ จะเป็น 1NF ก็ต่อเมื่อ ทุก ๆ โดเมนเป็นค่าครบ หน่วย (atomic values)
 - 🗕 ค่าของ Attribute เป็นค่าครบหน่วย
 - ตรวจดูว่า 1 Attribute มาจาก 1 โดเมน
 - Attribute (ชื่อและชนิดของ Attribute) ถูกต้องตามมาตรฐาน
 - ค่าของ Attribute สอดคล้องตามมาตรฐาน
 - ทุก Attribute ในแต่ละ record จะเป็น single value ไม่มีค่าของกลุ่ม
 ข้อมูลที่ซ้ำกัน (Repeating Group)
 - ข้อมูลทุกแถว (Tuple) ต้องมีค่าไม่ซ้ำกัน



ตารางที่มีลักษณะข้อมูลเป็น Repeating Group

s_id	s_name	s_surname	c_id
542107030005	กฤษดา	อิ่มรุ่งเรื่อง	CE354
			CE498
			CE305
554607030015	สาวิตรี	วิไลโรจน์	CE354
			CE356
			CE215
			CE327



First Normal Form: 1NF

X	<u>s_id</u>	s_name	s_surname	<u>c_id</u>
	542107030005	กฤษดา	อิ่มรุ่งเรื่อง	CE354
	542107030005	กฤษดา	อิ่มรุ่งเรื่อง	CE498
7	542107030005	กฤษดา	อิ่มรุ่งเรื่อง	CE305
	554607030015	สาวิตรี	วิไลโรจน์	CE354
	554607030015	สาวิตรี	วิไลโรจน์	CE356
+	554607030015	สาวิตรี	วิไลโรจน์	CE215
	554607030015	สาวิตรี	วิไลโรจน์	CE327



First Normal Form: 1NF

- ดูตารางความสัมพันธ์ PRODUCT เป็นรูปแบบบรรทัดฐานที่หนึ่ง หรือไม่
- ถ้าไม่เป็นรูปแบบบรรทัดฐานที่หนึ่ง ให้ระบุส่วนที่ทำให้ไม่เป็น และทำ ให้เป็นรูปแบบบรรทัดฐานที่หนึ่ง

PRODUCT

<u>PID</u> PNAME		PRICE
P001	Red shirt size 38	199
P 002	Blue shirt size 38	199
P003	Red shirt size 40	199
P_004	Blue shirt size 40	199
P005	Black Jeans	399



การพึ่งพิงเชิงฟังก์ชัน (Functional dependencies: FDs)

- X -> Y อ่านว่า "X functionally determines Y" หรือ "X arrow Y"
 - X เป็นตัวกำหนด (determinant)
 - Y เป็นตัวพึ่งพิง (dependent)
- ฉิยาม
 - ให้ r เป็นความสัมพันธ์
 - X และ Y เป็นเซตย่อยของกลุ่ม Attribute ใน r
 - X -> Y ก็ต่อเมื่อ แต่ละค่าของ X มีค่า Y หนึ่งค่าในความสัมพันธ์ r
 - หรือ ถ้ามี 2 ทูเพิลใน r มีค่า X เหมือนกัน จะต้องมีค่า Y ค่าเดียวกัน
 - X เป็นตัวกำหนดค่า (determine) ของ Y หรือค่าของ Y ขึ้นอยู่กับค่าของ X (Y depends on X หรือ Y เป็น FD กับ X)



การพึ่งพิงเชิงฟังก์ชัน (Functional dependencies: FDs)

• แบ่งได้เป็น 2 รูปแบบ คือ

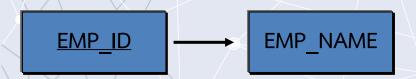
— แบบ 1: 1

ค่าของ Attribute ทางด้านซ้าย 1 ค่า

บ่งบอกค่าของ Attribute ทางขวา 1 ค่า

— แบบ m: 1

ค่าของ Attribute ทางด้านซ้ายตั้งแต่ 1 ค่าขึ้นไป บ่งบอกค่าของ Attribute ทางขวา 1 ค่า



เมื่อระบุค่าของรหัสพนักงาน (EMP_ID) จะ สามารถทราบชื่อพนักงาน (EMP_NAME) ที่มี รหัสตามที่ระบุได้



Full Functional Dependencies (Full FDs)

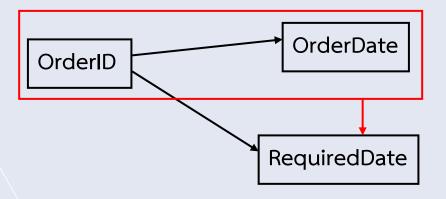
- การพึ่งพิงเชิงฟังก์ชันจะเป็น<u>การพึ่งพิงเชิงฟังก์ชันแบบเต็ม(full functional</u> dependencies: full FDs) ก็ต่อเมื่อค่า Y ขึ้นกับค่า X เท่านั้น
- หรือพูดอีกอย่างว่า ค่า Y ไม่ขึ้นกับเซตย่อยของ X
- ถ้าไม่เป็นไปตามเงื่อนไขนี้ เรียกว่าเป็น<u>การพึ่งพิงเชิงฟังก์ชันบางส่วน(partial dependencies)</u>



แผนภาพเอฟดี (FD Diagram)

OrderID	OrderDate	RequiredDate	CompanyName	Address	City	PostalCode
10248	04-Aug-94	01-Sep-94	Vins et alcools Chevalier	59 rue de l'Abbaye	Reims	51100
10249	05-Aug-94	16-Sep-94	Toms Spezialitäten	Luisenstr. 48	Münster	44087
10250	08-Aug-94	05-Sep-94	Hanari Carnes	Rua do Paço, 67	Rio de Janeiro	05454-876
10251	08-Aug-94	05-Sep-94	Victuailles en stock	2, rue du Commerce	Lyon	69004
10252	09-Aug-94	06-Sep-94	Suprêmes délices	Boulevard Tirou, 255	Charleroi	B-6000

• จากความสัมพันธ์ที่ให้มา สามารถเขียนแผนภาพเอฟดีได้ดังนี้



- FD1: OrderID, OrderDate -> RequiredDate
- FD2: OrderID -> OrderDate
- FD3: OrderID -> RequiredDate



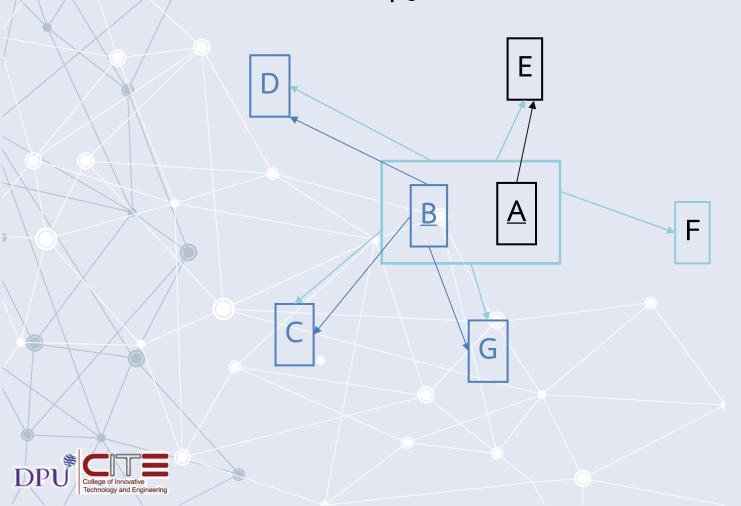
รูปแบบบรรทัดฐานที่สอง (Second Normal Form: 2NF)

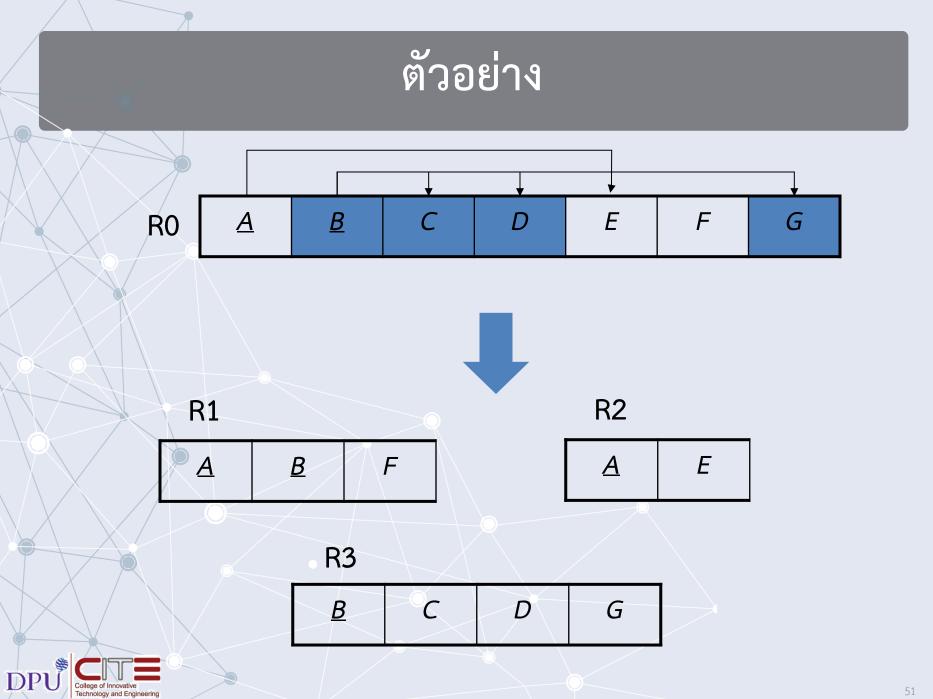
- ความสัมพันธ์ใด ๆ จะเป็น 2NF ก็ต่อเมื่อ
 - ความสัมพันธ์นั้นเป็น 1NF และ
 - Attribute ที่ไม่ใช่กุญแจ (non-key attribute) มีการพึ่งพิงเชิงฟังก์ชันแบบ เต็มกับกุญแจหลัก
 - ต้องไม่มี Partial Dependency (การขึ้นต่อกันบางส่วน)

• สรุปก็คือ Second Normal form : 2NF เป็นการขจัด Attribute ที่ ไม่ขึ้นกับทั้งส่วนของคีย์หลักออกไป เพื่อให้ Attribute อื่นทั้งหมดขึ้น ตรงกับส่วนที่เป็นคีย์หลักทั้งหมดเท่านั้น



• กำหนดให้ A และ B เป็นกุญแจหลัก





• จาก Relational Schema ที่กำหนดให้ จงเขียนแผนภาพ FD และ ตอบว่า เป็น 2NF หรือไม่ เพราะเหตุใด

	<u>SSN</u>	<u>PNUMBER</u>	HOURS	ENAME	PNAME	PLOCATION



การแปลงรีเลชันรูปแบบบรรทัดฐานระดับที่ 1 ให้เป็นรีเลชันที่มีรูปแบบ บรรทัดระดับที่ 2

รรายงานการทำงาน

รหัย	1 ව්ට	ประเภท	อัตรา	จำนวน	วหัสผู้	อัตราค่า					
ดนงา	น ดนงาน	ดวามซ้ำนาญ	โบนัส	ช.ม.ชั้นดำ	ดวบดูม	แรง/ช.ม	รหัสสถาน	ที่อยู่สถาน	ประเภทสถาน	วันที่เ ริ่ ม	จำนวน
							ที่ก่อสร้าง	ที่ก่อสร้าง	ที่ก่อสร้าง	ทำงาน	ช.มที่ทำ
			\sim					•			
									\simeq		
ดนง	คนงาน-ชำนาญ										
<u>รหัต</u>	1 20	ประเภทจ	าวามอัตร	า จำนวน	วหัสผั	อัตราค่า	1				
ดนงา				สช.ม.ขันตำ			1				
1245						~~~	-				
1 4245	100 0 03	150 M 1 1 1 1 1 1 1 1 1 1 1 1	1 3.50	40	1411	16.0	1				

การทำงาน-ก่อสร้าง

<u>รหัส</u> <u>ดนงาน</u>	รหัสสถาน ที่ก่อสร้าง	ที่อยู่สถาน ที่ก่อสร้าง	ประเภทสถาน ที่ก่อสร้าง	วันที่เริ่ม ทำงาน	จำนวน ช.มที่ทำ
1245	516	128 ดาด นร ้าว 78	บ้านพักอาจัย	01/01/49	50
1245	311	25 ดีตม	สำนักงาน	08/01/49	40
2521	516	128 ดาดพร้า ว 78	บ้านพักอาจัย	01/01/49	56
2521	450	18 พหลโยธิน 55	ร้านด้า	15/01/49	45
2521	431	50 อินทามระ 27	บ้านพักอาจัย	23/01/49	36



พิเธฏ บากบี

ปูน

2.50

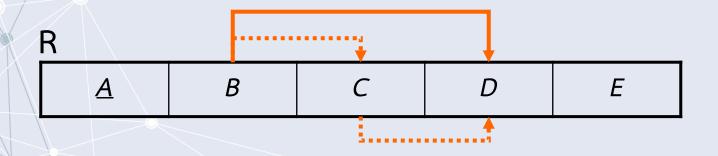
• จงเขียนแผนภาพเอฟดีของตารางความสัมพันธ์ PROJECT_EMPLOYEE และทำให้เป็น 2NF

PROJECT_EMPLOYEE

	<u>P#</u>	PNAME	<u>E#</u>	ENAME	JOB	CHARGE	HOURS
	P01	Web	001	Tommy	Programmer	300	48
	P01	Web	002	Рорру	SA	500	32
S	P02	Database	003	Jessie	DBA	400	40
	P03	Pocket PC	001	Tommy	SA	500	50
	P04	Web	002	Poppy	Programmer	300	20
	P04	Web	004	Tommy	DBA	400	36
>	P04	Web	001	Tommy	Programmer	300	30
	P05	Mobile App	002	Poppy	Programmer	300	25
PĮ	•••	•••	•••	•••	•••	•••	•••

Transitive Functional Dependencies (Transitive FDs)

• ในความสัมพันธ์ R(A, B, C, D, E)



- ถ้า B -> C และ C -> D
- จะได้ว่า...
- B -> D เป็น Transitive FD



1	SID	SNAME	DEPARTMENT	FACULTY
	51024	นายอภิชัย พุ่มศรี	วิศวกรรมคอมพิวเตอร์	วิศวกรรมศาสตร์
	54041	นายเศรษฐพงษ์ อิ่มสุวรรณ์	วิศวกรรมคอมพิวเตอร์	วิศวกรรมศาสตร์
	54004	นางสาวพิมพ์สร ว่องวิการณ์	วิศวกรรมดิจิทัลมีเดียและระบบเกม	วิศวกรรมศาสตร์
	53007	นางสาวปรารถนา ทรัพย์ประดิษฐ์	วิศวกรรมโทรคมนาคม	วิศวกรรมศาสตร์
				•••



ต่อไปนี้เป็น FDs หรือไม่?

- SID → SNAME
- SID → DEPARTMENT
- SID → FACULTY
- SNAME → DEPARTMENT
- SNAME → FACULTY
- DEPARTMENT → FACULTY
- FACULTY → DEPARTMENT
- FACULTY → SID











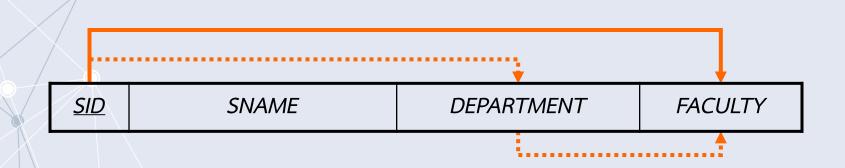








Transitive FDs



- SID → SNAME
- SID → DEPARTMENT
- SID → FACULTY
- DEPARTMENT → FACULTY



รูปแบบบรรทัดฐานที่สาม (Third Normal Form: 3NF)

- ความสัมพันธ์ใด ๆ จะเป็น 3NF ก็ต่อเมื่อ
 - ความสัมพันธ์นั้นเป็น 2NF <u>และ</u>
 - —ไม่มี <u>Transitive FD</u> ระหว่าง <u>primary key กับ non-key attribute</u>
 - Attribute ที่ไม่ใช่ Primary key ต้องไม่ขึ้นต่อกันเอง



คำถาม...

ความสัมพันธ์นี้เป็น 3NF หรือไม่

SID SNAME DEPARTMENT FACULTY





คำถาม...

• จงทำให้เป็น 3NF

	SID	SNAME	DEPARTMENT	FACULTY
7				



ເฉลย...

• ต้องกำจัด Transitive FD ระหว่าง primary key กับ non-key attribute



Boyce/Codd Normal Form: BCNF

- ความสัมพันธ์ใด ๆ จะเป็น BCNF ก็ต่อเมื่อ
 - ทุก ๆ determinant ต้องเป็น candidate key



A B C D E F	G
-------------	---

- Candidate key = A, B, C
- Full FDs

$$-B \rightarrow D, E$$

$$-C \rightarrow D, E$$

$$\rightarrow$$
 E \rightarrow F, G

$$-F \rightarrow G$$





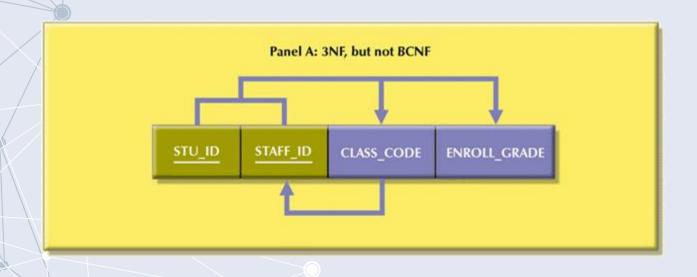


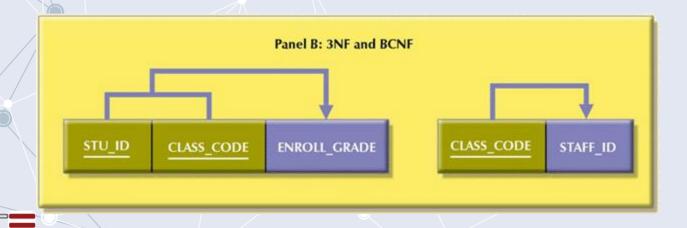


A B C D E

E F G







Technology and Engineering

SID	CID	CNAME	GRADE
54001	CE351	การออกแบบฐานข้อมูล	Α
54001	CE341	วิศวกรรมซอฟต์แวร์	C+
55002	CE351	การออกแบบฐานข้อมูล	C+
55002	CE111	คอมพิวเตอร์เบื้องต้น	D
52003	CE351	การออกแบบฐานข้อมูล	В
52003	CE111	คอมพิวเตอร์เบื้องต้น	Α



• จงพิจารณาว่า จากตัวอย่างในหน้าที่แล้ว เป็น 1NF 2NF 3NF และ BCNF หรือไม่ หากไม่เป็นจงแก้ไขให้เป็น





Questions & Answers